New Evaluation Techniques
For Estimating the Impacts of Science and Technology Innovations

Paper 1

A Credible Approach to Benefit-Cost Evaluation for Federal Energy Technology Programs

Gretchen B Jordan, Sandia National Laboratories, gbjorda@sandia.gov
Rosalie Ruegg, TIA Consulting, Inc., ruegg@ec.rr.com

Lawmakers want to know the returns on their investment. The difficulties in measuring benefits and costs, and attributing some portion of those to specific research and technology development (R&D) programs, has meant that many such studies are not seen as credible by those lawmakers. This paper describes a methodology that improves upon an already credible approach developed for a 2001 National Research Council study: "*Energy Research at DOE: Was It worth It*?" Three benefit-cost studies using this modified approach will be completed by the U.S. Department of Energy's Energy Efficiency and Renewable Energy in 2009.

The key evaluation questions to be answered in these studies are:

- To what extent have programs produced "actual" economic benefits (energy-savings, renewable market growth, and other positive economic effects) relative to the next best alternative?
- To what extent have programs promoted environmental benefits and enhanced energy security by providing alternative energy sources and energy efficiency, and protecting existing sources?
- Would today's commercialized technologies have happened at the same time, with the same scope and scale, and with the same extent of deployment without DOE involvement?
- To what extent do benefits attributable to DOE involvement exceed DOE R&D expenditures?

The modified methodology will implement the retrospective benefits estimation recommendations of the May 2002 *EERE Strategic Technical Review* and the *DOE Benefits Workshop* held in 2003.

Economic performance metrics that are calculated are Net benefits, Benefit-cost ratio, and Internal Rate of Return. All of these are social & public returns that include private returns but these are not calculated separately. The types of benefits and costs included in calculation of metrics are Investment costs, Energy costs, Other operating costs, Maintenance, repair, and replacement costs, and Rents or royalty payments.

Benefits and costs for selected technology "winners" are calculated compared against the next best alternative. Differences are expressed as year-by-year dollar cash flows, adjusted to constant dollars, and amounts are "discounted" for the time value of money (apart from inflation) using OMB-specified real discount rate. Additionally, a "Cluster approach" is used that compares benefits of larger elements of a program to investment costs of the entire program. The cluster approach describes the larger program and government's role, as well as the specific technology elements. It generates a minimum estimate of return for the program without performing detailed analysis of everything in it, thereby increasing feasibility of the analysis.

Environmental and Security benefits are also assessed. Environmental benefits quantified will focus on reduction in air pollution. Security benefits will focus on cumulative fuel savings (oil and natural gas), expressed in physical terms. Important benefits that cannot be measured quantitatively within the scope of a study are discussed qualitatively.

Knowledge Benefits were added to the NRC framework by experts attending the 2003 DOE Benefits Conference. This EERE benefit-cost methodology calls for a more comprehensive and quantitative assessment of knowledge benefits of the technologies selected for study than have previous benefit-cost studies. Studies identify and document linkages between more than three decades of DOE R&D and, in the 2009 studies, commercial renewable power generation. Studies use interviews with scientists, engineers, and administrators in government, as well as manufacturers and their customers. They also review documents and databases for evidence of linkages, and apply the tools of patent and publication citation analysis, and publication co-author analysis to shed further light on the paths through which outputs of DOE-funded R&D have been disseminated to producers and users.

In contrast to the 2001 NRC study, the modified approach requires a case-by-case assessment of additionality—the difference that DOE made in the outcome. Rather than use a rule of thumb, the EERE modified NRC approach will examine an array of ways additionality can occur. These include acceleration in the development and introduction into the market of a technology, change in the scale, performance characteristics, function, and cost of a technology (e.g., its emissions profile or its durability), acceleration in the rate at which a technology gains share of the existing domestic market after it is initially introduced, increase in access of U.S. firms to foreign markets, expansion in the total size of domestic and foreign markets, and reduction in the total investment costs (combined public and private) of achieving the desired outcome as compared with achieving the same with industry going it alone.

To have more comparability across EERE benefit-cost studies, this approach will use generic logic models to describe the changes over the time period in technologies and markets using a common language. Activities, external influences, and progress over time in R&D will use the stages of the Stage Gate process: Preliminary investigation, detailed investigation, development, validation, and commercial launch. Activities, external influences, and progress over time in the "readiness" of the market and technology adoption will use four market domains. In addition to the usual domain of the end user that looks at adoption of a technology, there are three market infrastructure domains: Information, Policy/Government, and Business. Examples of activities in these infrastructure domains are the Wind Resources database, EERE efforts on new appliance standards, and partnerships with industry on solar manufacturing R&D.

Paper 2

Techniques for Evaluating New Scientific Instruments

Jonathon Mote, University of Maryland, jmote@socy.umd.edu
Aleia Clark, University of Maryland, alclark@socy.umd.edu

This paper proposes an evaluation technique for evaluating the impacts of new scientific instruments in a way that avoids the pitfalls of “economic-only” cost-benefit analysis and meets the needs of the customer organization and Congress. On several occasions, Congress has rejected the requested appropriation of 700 million dollars to place a new suite instruments, the Hyperspectral Sounder (HES), on the GOES T, a new weather satellite to be launched in 2016, by the National Oceanographic and Atmospheric Agency (NOAA). Part of the reason for the rejection is that the economic cost-benefit analysis completed by NOAA indicates only marginal gains from this investment. At the macro level of predicting the national weather, this is true. Over the last twenty years there has been considerable gain in the accuracy of the sixth and seventh day weather predictions. Of course, you feel quite differently about this at the local or meso (regional level) because you focus on the many errors--the unannounced snowstorm that shuts down your city for three days, the unexpected tornado that kills people in your state, or the sudden rainstorm that drowned your picnic. The specific strength of this new set of instruments is that it can be more specific about where severe weather will occur.

Another reason why Congress has been reluctant to move ahead, outside of the enormous budget deficits, is the lack of support of the National Weather Service (NWS) within NOAA. Again, there are some understandable reasons for this resistance. NWS uses a global model from which meso models are derived. In fact, the local weather predictions are made by various meteorologists who work for the radio and television stations in your community. The macro models presently push the limits of super computers and cannot easily absorb real time data over short periods (say six hours or less). To do so, requires totally reworking the entire set of complex equations used for weather prediction involving inputs from around the world.   This is a very expensive and time consuming task and it is not clear that there is a large enough computer to handle this new information. This is a classic example of path dependency because of an existing technology and cognitive model along with considerable sunk costs in the present system.

Therefore, our new evaluation techniques have to accomplish three objectives:

1) Develop new criteria that avoid the pitfalls of the economic cost-benefit analyses
2) Create a case that would convince Congress despite budget deficits that investment in the HES is valuable for their constituents
3) Propose a political strategy that overcomes the resistance of NWS

In addition, we want our discussion of these new evaluation techniques to be in terms that allow others to apply these ideas to situations other than weather forecasting.

As indicted above, the strong point of HES is that it predicts sudden weather changes (within six hours) on a meso scale (region or even locality). It can do this because it is measuring the following weather parameters: vertical moisture profiles, vertical temperature profiles, derived stability indices, derived motion of winds, moisture flux and ozone total. Together these measure sudden inversions, the major cause of unexpected instabilities of weather associated with tornados, severe winter storms, floods, and the like.

Therefore, the solution to the first problem is to focus on improvements in warning time for each of these severe weather events. To take one example, to move from 13 minutes to 13 hours of warning time for a tornado saves lives. Note, that warning time does not reduce the destruction of property. This is still there. But this is one reason why economic benefits are largely but not always beside the point and especially for extreme weather events. The damage still exists although it might be slightly mitigated but one can save lives. Besides tornados, thunderstorms including flooding, and winter storms, we also want to focus on sudden decreases in air quality and its health consequences.

The political solution for Congress is to chart every unanticipated severe weather event in each congressional district and state during the past year to indicate how many of their constituents would be affected by earlier warning times. Of course, the question is what evidence can be marshaled for a system that is not operational, one of the major parts of our presentation.

The strategy for handling the NWS is to advocate the creation of separate meso models, which can be much simpler and adapted to the specific geographical regions of the United States and of course constructed to absorb real time data. In other words, the argument is to supplement and not to argue that the NWS should change its modeling.

Paper 3

A New Evaluation Strategy for Measuring the Returns on Investments in Medical Research

Jerald Hage
University of Maryland, Hage@socy.umd.edu

As part of the continued efforts of the Center for Innovation to develop new evaluation strategies, we developed a new framework and proposed this in an invited white paper for a major assessment called for by the Canadian Academy of Health Sciences. What makes this framework new is that is focuses on the treatment sector because this is where the greatest variations in both research investments and potential returns are concentrated. By focusing on the treatment process, a very fine-grained evaluation is possible, which then can be aggregated to the macro level, the level of policy makers. But since it starts with the treatment process, it *can pinpoint where either too much money in research has been invested or not enough.* This is a very important question for policy-makers.

One critical component of this framework is that it provides precise definitions of what is the treatment process as distinct from the research process, the differences in where treatments are provided and where research is conducted, and the distinctions between the micro, meso, and macro levels of evaluation. These definitions are important because of the mismatch among them.

The specific metrics of the framework are:

1. Metrics of health care impact by stage in the treatment process;

2. Metrics of research investment by arenas within the production of medical knowledge within the specific treatment sector;
3. Metrics of contributions to scientific knowledge;
4. Metrics of network gaps in the production of innovative treatment protocols;
5. Metrics of economic and social benefits of medical research.

The key starting point is the treatment process, which is defined by the differences in the nature of the illness, injury or health care problem that is being treated. The metrics for this are groups according to the four stages: prevention, intake and assessment, treatment, and post-treatment including long term care. Carefully specifying the stages in the treatment process associated with a particular morbidity allows for a fine-grained set of health care impact metrics or indicators. One could make additional distinctions within these four stages. For example, one might want to distinguish between diagnosis and prognosis. In addition to the four stages of the treatment process, we have added a category, knowledge about the health care problem, because a major part of biomedical and population research focuses on the development of understanding about the health care problem that eventually can lead to either prevention or treatment. Two to three indicators are suggested for each stage. The problem of actual vs. potential benefits--an issue that plagues many evaluations--is also discussed.

To place the evaluation in its proper context, the kinds of investments made in medical research, both human and capital, are classified according to the specific stages of the treatment process. This highlights gaps. Another set of measures deals with detecting gaps in the idea innovation network (Hage and Hollingsworth, 2000) associated with a specific treatment sector. This is particularly important given the presence of a valley of death between medical research and the development of industrial innovations perceived to exist, again illustrating the advantages of selecting the sector level in the health care delivery system.

A special section on metrics for knowledge contributions is suggested as well, given the importance of this for most academics. In this, a special emphasis is placed on the international impact of these contributions.

Metrics for economic benefits flow naturally from the specific indicators for each stages of the treatment process. Examples include value of illness days saved from decline in morbidity incidence, reduction in the costs of tests for diagnosis, reduction in the patient's costs of waiting, value of days saved in hospitalization, value of days saved in rehabilitation and after care, etc. Surprisingly, focusing on the stages in the treatment process, which would seem to involve more work, simplifies the task of specifying the specific benefits of a particular kind of research finding. Finally a number of societal benefits are indicated as well including such things as increased equality in health care and duration of life by class and gender.